Assessment of uncertainty in functional–structural plant models

E. David Ford* and Maureen C. Kennedy
School of Forest Resources, University of Washington, Seattle, WA 98195-2100, USA
*For correspondence. E-mail edford@u.washington.edu

Received: 24 November 2010  Returned for revision: 11 February 2011  Accepted: 10 March 2011  Published electronically: 18 May 2011

INTRODUCTION

The motivation for developing and using functional–structural plant models (FSPMs) is to construct theories about the integration of form and function in plants and their interaction with the environment (Godin and Sinoquet, 2005). FSPMs generally comprise many component functions, and bounding decisions are made about what processes should be included and the detail with which they should be represented. These choices introduce uncertainty into the model because we cannot be certain that all of the functions in an FSPM have been selected, represented and connected effectively. Indeed, we construct FSPMs to examine just how adequate the abstraction represented by the model might be. Nor can we always be certain that an FSPM predicts all the features of a system that it should — some features may be discounted. For example, Fourcaud et al. (2008) point out that many growth models are limited by source activity but suggest that in natural conditions sinks will more often drive growth than source activity because sink activity is often controlled by finite soil resources or development conditions.

The use of FSPMs to explore theories requires that they should be assessed in a way that will define their uncertainty and lead to improvements. We define the types of uncertainty found in FSPMs. We present a method of assessment that takes account of their complexity, particularly that they tend to estimate quantities of different types — not only dimensions such as height and weight but conditions such as nutrient concentrations or presence or absence of particular structures. We anticipate that in general model assessment will be progressive where particular components of a model may be improved or replaced. Application of the method is illustrated by development of a probabilistic architectural model for branch growth in old-growth conifers into one defining its causal basis.

MATERIALS AND METHODS

Components of uncertainty

Decisions made when constructing an FSPM about what should be included in the model and what it should be able to predict lead to two important categories of model uncertainty: non-uniqueness and incomplete specification (Turley and Ford, 2009). Non-uniqueness is the characteristic that there may be alternative representations of some component of a model and no a priori way of deciding between them. Frequently, an objective for constructing an FSPM is to decide which representation of a number of associated processes is most effective. Incomplete specification is the characteristic that a model may explain some, but not all, of the...
observed features of a system. We generally require that a model attains its objectives in making predictions about some feature without producing a distorted prediction of other features. To define these categories of uncertainty we need to make simultaneous assessment of multiple criteria.

Model uncertainty should be defined and calculated progressively as improvements and comparisons are made between successive versions. Following Turley and Ford (2009) we can say:

1. Uncertainty is reduced as the model can simultaneously simulate an increasing number of assessment criteria. These should be selected to test specific aspects of the theory being investigated and acceptable limits must be set for them. This reduces incomplete specification.
2. Uncertainty is reduced as the distribution of values for parameters simulating multiple assessment criteria within acceptable limits becomes unimodal and with reduced range. This reduces non-uniqueness in the model. The ideal is that one set of parameters for a model should be able to simulate all of the required assessment criteria. Frequently, different sets of parameter values may be required to simulate particular combinations of assessment criteria which indicate uncertainty in the model and may require change of additional functions.
3. Uncertainty is reduced as the acceptable limits for assessment criteria are defined with increasing precision. This links uncertainty in the model to empirical research used in its construction.

**Pareto Optimality**

The objective is to discover if the model can achieve multiple assessment criteria simultaneously with distributions of parameter values that are consistent with the biological processes being studied. As more criteria are added to an assessment it can be anticipated that the model will be more effective in some predictions than others but may not achieve all criteria, at least not with a unique set of parameter values. The model’s effectiveness in achieving multiple assessment criteria can be examined by calculating its Pareto Optimality (Fig. 1).

A set of parameter values, one for each parameter in a model simulation, is a model parameterization. The Pareto Optimal Set is the collection of parameterizations that achieve unique combinations of the selected assessment criteria within the limits specified for them.

Consider just two assessment criteria, A and B, for both of which there are measured data. Comparisons of differences between measured and simulated values, |measured−simulated|, can be graphically represented (Fig. 1A). If minima for |measured−simulated| can be found for both A and B at the same time then there would be a single point closest to [0, 0]. In the case illustrated the minimum for |measured−simulated| for variable A is not associated with the minimum |measured−simulated| for variable B. For continuous variables we are likely to decide on some allowable magnitude of |measured−simulated| (Fig. 1A) but no parameterization may produce |measured−simulated| values that are less than the allowable magnitude for both A and B (Fig. 1A). Where only a few assessment criteria are of interest then a series of pairwise plots similar to Fig. 1(A) can be informative about the model’s ability to simulate different variables effectively at the same time.

Using allowable magnitudes for |measured−simulated| for each assessment criterion enables a binary classification of success or failure to be made for each parameterization of the model and individual parameterizations can then be ranked by their relative performance across all of the criteria. Typically, we expect that individuals will achieve some but not all criteria. In Fig. 1(B) results illustrating possible outcomes for five parameterizations are shown. Parameterizations 1, 2 and 3 each achieve unique combinations of assessment criteria. Parameterizations 4 and 5 achieve some assessment criteria but do not do as well as at least one other parameterization and are said to be dominated. The non-dominated parameterizations, 1, 2 and 3, constitute the Pareto Optimal Set.

![Fig. 1.](https://academic.oup.com/aob/article-abstract/108/6/1043/209748) Fig. 1. (A) Projected results from an attempt to minimize |measured−simulated| simultaneously for two assessment criteria, A and B. Results for individual parameterizations of a model are represented by the points. Minimization for both A and B are not achieved simultaneously. Low values in one are associated with high values in the other. Taken together these results can be used to define a Pareto Frontier, which would run through the most concave line of points. Where a minimum required |measured−simulated| is specified, represented by a and b then parameterizations can be classed as acceptable or not for that particular assessment criterion. (B) Projected results for five parameterizations (1–5) tested against four assessment criteria (I–IV) for which minimum values of |measured−simulated| have been specified and used to classify the parameterization as successful (filled) or unsuccessful (clear). Parameterizations 1, 2 and 3 each achieve unique combinations of assessment criteria, i.e. they are non-dominated. Parameterizations 4 and 5 do achieve some assessment criteria but they are dominated by parameterizations that achieve more assessment criteria: parameterizations 4 and 5 are both dominated by parameterization 3.
Calculation of the Pareto Optimal Set

Calculation requires an exploration of the feasible parameter space. For a large parameter space a full exploration can be computationally prohibitive so we use an evolutionary algorithm to converge to an approximation of the Pareto Optimal Set (Coello Coello et al., 2002). Evolutionary computation (Reynolds and Ford, 1999; Komuro et al., 2006) first establishes a population (say 100) of parameterizations for the model. The multiple outputs being used in model assessment are calculated for each individual parameterization of the population. The highest ranked individuals, those that are non-dominated, are retained and used to breed new individuals for a new population, the second generation, which in turn is tested in the same way. ‘Breeding’ is carried out either by simple exchange of parameter values between individuals, which explores the effect of different combinations, or by small random changes to parameters (Komuro et al., 2006). This is repeated for a number of generations (say 500) until a suitable approximation of the Pareto Optimal Set or frontier is achieved.

Procedure for model analysis

Calculation of the Pareto Optimal Set can be applied during the development and progressive exploration of the functioning of a model (Fig. 2). This procedure, the Pareto Optimal Assessment Cycle (POMAC), adapted from Reynolds and Ford (1999), involves a number of stages each of which can provide information about the model.

Stage 1. Initial definition of the model and its objectives

(a) Define the model and what it should predict.
(b) Define the parameters in the model and ranges within which values can be sought.
(c) Define criteria to be used in assessment of model functioning and their acceptability ranges. Assessment criteria are likely to include what the model should predict but should also include criteria that inform about the effectiveness of model functioning. Turley and Ford (2009) suggest three categories of assessment criteria should be used: (1) tests of the quantitative domains for model outputs based on empirical data; (2) examination of features explicitly predicted by the model; and (3) tests between alternative structures in model formulation.

Stage 2. Evolutionary computation

(a) Calculate the Pareto Optimal Set. Technical details are discussed in Komuro et al. (2006) and Turley and Ford (2009).
(b) Examine the adequacy of parameter search ranges. For practical purposes of restricting computation, time bounds may be set for permissible parameter values and it is important to consider if these have been restrictive; for example, if it is found that parameter values are clustered close to one of the bounds then either the bounds need to be changed and/or the cause of the clustering explored.
(c) Examine if there are deficiencies in the mathematical structure of the model. At this stage apparently unusual or unexpected results may be found that can indicate problems with the structure of the model and particularly mathematical representation in functions (for examples see Reynolds and Ford, 1999).

Stage 3. Examination of the Pareto Optimal Set

(a) Examine the distribution of parameter values for members of the Pareto Optimal Set. This can be a complex and time-consuming process and it is often helpful to reduce the number of dimensions to examine. This is achieved by clustering the parameter values and reducing the number of dimensions to two or three. This can be done by using techniques such as principal component analysis, which reduces the number of dimensions to two or three while retaining as much of the variance as possible.
(b) Examine the structure of the Pareto Optimal Set. This involves looking for patterns in the parameter values that are produced by the model. This can be done by using techniques such as cluster analysis, which groups similar parameter values together and identifies the patterns that are produced by the model.

Stage 4. Look for mathematical deficiency

(a) Look for mathematical deficiency in the model. This involves looking for mathematical deficiencies in the model that may be causing the model to produce unexpected results. This can be done by using techniques such as sensitivity analysis, which identifies the parameters that are most sensitive to changes in the model and identifies the mathematical functions that are most sensitive to changes in the parameters.

Fig. 2. The Pareto Optimal Model Assessment Cycle. The main flow of the cycle is indicated by solid lines and filled arrowheads; revision processes are indicated by broken lines and open arrowheads. Following model construction and definition of what it should predict (centre), ranges must be defined for parameters and assessment criteria selected with suitable ranges. Following the first calculation of a Pareto Optimal Set the adequacy of parameter ranges must be examined, changed if necessary and the Pareto Optimal Set recalculated. Deficiencies in mathematical structure can lead to model revision. Both the structure of the Pareto Optimal Set and the parameter values producing it must be examined. Changing parameter ranges and/or assessment criteria can lead to an improved model. [Adapted from Reynolds and Ford (1999).]
Stage 3. Analysis of Pareto Set

(a) Examine the structure of the Pareto Optimal Set. The critical consideration is how many non-dominated groups comprise the set and investigating what determines this. Two techniques can be useful in this investigation: inclusion of different variables as assessment criteria and changing their acceptability ranges. Typically, at least some assessment criteria are based on empirical data but usually FSPMs can be assessed against some theoretical criteria as well. These may be simple but can be used to ensure a model is simulating within acceptable norms. For example, models simulating complete aerial branching structures might have assessment criteria for plant height and crown width, or their ratio. These may not have been measured but a range for values is likely to be known. Changing acceptability ranges for assessment criteria can reveal sensitivity of the model to requirements for predicting different criteria. With sufficiently wide ranges it may be possible for the model to achieve a Pareto Optimal Set that simultaneously includes all criteria. Progressive reduction can indicate which criteria are most difficult for the model to achieve (for examples see Komuro et al., 2006).

(b) Investigate the distribution of parameter values between and within parameterizations of the different member groups of the Pareto Optimal Set. Examination of the combination of parameter values that achieve different members of the Pareto Optimal Set can be instructive in revealing the dynamics of the model. For example, when two assessment criteria cannot be achieved simultaneously and clearly different values of a parameter are required to achieve each of them, then function(s) using that parameter may need to be modified. Distributions of parameter values that are bimodal or widely spread should be analysed and an examination of the model made to see if there are parameters that compensate for each other.

Stage 4. Define model uncertainty. This should be based on a description of the model. It should include description of the Pareto Optimal Set for a given set of assessment criteria with defined ranges and how achieving or failing these criteria contributes to understanding of model function. There should be analysis of parameter values attaining different groups of the Pareto Optimal Set. Definition of uncertainty should include description of the adequacy of the selected assessment criteria.

RESULTS

To illustrate the definition and calculation of uncertainty we use an example of a model for branch growth. In an old-growth forest in south-west Washington State, USA, 400-year-old trees of Douglas-fir ( _Pseudotsuga menziesii_ ) had stopped height growth (Ishii et al., 2007) and lateral crown expansion (Ishii, 2000). However, from a population dynamics study of the same stand using repeated measures made over four decades, Franklin and DeBell (1988) predict that if current mortality rates continue Douglas-fir trees could persist for a further 755 years. This raises the question of how foliage and crowns are maintained in these old-growth trees.

Ishii and Ford (2001) report that shoots and foliage on branches of these old-growth Douglas-fir are constantly renewed by non-damage-related epicormic shoot production, which results in reiteration of shoot cluster units (SCUs), an architectural unit of shoot organization within branches (Fig. 3). Phases of SCU development were identified from initiation through maximum shoot and foliage production to decline and failure to produce new shoots. SCUs exist for between 12 and 20 years depending on tree size and position of branch in the crown. Up to seven generations of SCUs were identified, each one arising on top of an older one.

Ishii and Ford (2001) suggest that this non-damage-related epicormic regeneration of foliage maintains shoots and foliage of old-growth Douglas-fir after height growth and lateral crown expansion cease and may contribute to prolonging tree longevity. This is a phenomenological rather than a causal explanation (greater longevity is explained by the phenomenon of non-damage-related epicormic reiteration). The approach was first to develop a geometrically based probabilistic model, BRANCHPRO, defining the process of non-damage-related epicormic reiteration and then develop its causal explanatory power using multi-criteria assessment. We describe this development by applying the four stages of multi-criteria assessment (Fig. 2).

Stage 1. Initial definition of the model and its objectives.

(a) Define the model and what it should predict.

The starting point is a model for development of the architectural structure of old-growth branches of _P. menziesii_ (Kennedy et al., 2004; Kennedy and Ford, 2009). Architectural hierarchy is established, with shoots along the branch main axis assigned order 1, and for each lateral junction the order increases by 1 (Fig. 4). A single bud is suppressed for possible epicormic initiation for each order 1 shoot. Each year, for each actively proliferating shoot, a random number of daughter shoots is drawn from a probability distribution with some branching rate (r), with the number of daughter shoots restricted to values of [0; termination of shoot; 1; elongation of shoot axis but no laterals produced; 2; elongation of shoot axis and production of one lateral; 3; elongation of shoot...
Fig. 4. (A) The basic structure of sequential growth with a main axis continuing growth each year, producing two lateral axes and a suppressed bud that may grow into an epicormic shoot. Lateral axes may continue to grow. These bud developments are modified by a series of probabilities. If the main axis ceases to grow it may be replaced by the growth of a lateral axis. Lateral axes may cease growth and this affects the probability that a suppressed bud may develop into an epicormic shoot. (B) Suppressed buds may develop into epicormic shoots that in turn may grow sufficiently, depending on proximity to other SCUs and hydraulic path length or number of hydraulic junctions.

The objective for the next stage of model development was to define a possible causal explanation, particularly to test two major causes proposed to explain decline in ageing trees. One is associated with plant–water relationships and particularly increased resistance to water conduction as trees age (Ryan and Yoder, 1997; Hubbard et al., 1999, 2002; Bond, 2000; Woodruff et al., 2004). The other postulates that demands for carbon increase relative to the tree’s capacity to supply those demands and still produce new foliage (Grier and Logan, 1997; Cannell, 1989; Hunt et al., 1999). Kennedy et al. (2010) developed BRANCHPRO3 to examine if epicormic reiteration could possibly minimize resistance to water flow and/or carbon costs associated with increasing size.

Functions were developed relating the values of branching frequency (r) and epicormic bud release from suppression (p) to a set of independent variables (Table 1; Kennedy, 2008), with non-linear monotonic relationships (for details see Kennedy et al. 2010, appendix B). Constants k2j relate to the shape of those relationships:

\[
   r(\alpha, g, n_{over}, t_{base}, d_{base}) = r_0 + r_{ord} \left( 1 - \frac{\alpha}{\alpha} \right) + r_{gen} \left( \frac{g}{k_{24} + g} \right) + r_{over} \left( \frac{n_{over}}{k_{25} + n_{over}} \right) + r_{base} \left( \frac{t_{base}}{k_{26} + t_{base}} \right)
\]

\[
p(t_{bud}, a, d_{SCU}, d_{stem}) = p_0 + p_{bud} \left( \frac{t_{bud}}{k_{21} + t_{bud}} \right) + p_a * a + p_{SCU} \left( \frac{d_{SCU}}{k_{22} + d_{SCU}} \right) + p_{stem} \left( 1 - \frac{d_{stem}}{k_{23} + d_{stem}} \right)
\]

For each shoot functions are updated each year as values for independent variables change. Magnitude and direction (increasing or decreasing) of each component of the function that corresponds to a single independent variable (e.g. the bud age \( t_{bud} \)) are controlled by the value of the parameter associated with it (e.g. \( p_{bud} \)). A parameter value of zero means that the corresponding independent variable does not change the computed value of the function. A positive parameter value means the function value increases with the corresponding function component, and a negative value means the function value decreases with the corresponding function component. Each function also has an intercept \((p_0, r_0)\) providing a baseline function value not related to any independent variable. The value of branching frequency is calculated for
each foliated shoot with an active apical meristem, and the probability of epicormic bud release from suppression is calculated each simulation year for each suppressed bud. The independent variables represent possible biological features that may affect the value of either bifurcation or bud release from suppression (Table 1).

The function for branching frequency (eqn 1) contains a term for the number of parental epicormic generations (representing the number of constrictive turns in water conduction), the age of the epicormic bud at an SCU base when it was released (representing the likely greater resistance of the junction as buds age), the number of overlapping foliated shoots (foliage overlap is approximated by a count of the number of foliated shoots of the same generation or higher that intersect in the horizontal plane with the shoot being considered), the order of the branch and the generation of epicormic branching representing a gradual ageing.

(b) Define the parameters in the model and ranges within which values can be sought.

Each explanatory variable is associated with a parameter that determines the magnitude and direction of the effect of each variable on the probability that a bud is released to form a new epicormic shoot (Table 1). Each of these is associated with an individual parameter that defines the magnitude and direction of the effect of each variable on the mean bifurcation rate for a given actively proliferating shoot.

(c) Define criteria to be used in assessment of model functioning and their acceptability ranges.

The criteria used with BRANCHPRO3 fall into the three categories suggested by Turley and Ford (2009) and we describe how use of each contributes to explanatory power of the model and comment on limitations to criteria, themselves summarized in Table 2.

(i) Tests of the quantitative domains for model outputs. The following were used: empirical measures of branch length; number of foliated shoots; number of SCUs; and the branching frequency sequence used in defining dynamics of bud production and foliated shoot development.

The parameters determine the direction (positive or negative, i.e. decrease or increase in function value with increasing independent variable) and the magnitude (i.e. the size of the effect of increasing the independent variable on the function value) of the associated independent variable on the function value for each process [bud release from suppression (p) and bifurcation (r)]. Adapted following Kennedy et al. (2010).

### Table 1. Independent variables, associated parameters and underlying process assumptions in BRANCHPRO3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>Description</th>
<th>Process assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>p function</td>
<td>p₀</td>
<td>Baseline probability</td>
<td>If all other parameters or variables are zero, this is the probability of release. The effects of the variables are measured against this baseline.</td>
</tr>
<tr>
<td>titude</td>
<td>pₘₜₑₐₜ</td>
<td>Age of bud</td>
<td>The probability a bud is released is related to its time since suppression.</td>
</tr>
<tr>
<td>a</td>
<td>pₜ</td>
<td>Proportion of inactive axes</td>
<td>The inhibition of the bud declines as neighbouring growth slows, as measured by the proportion of axes immediate to the current bud (main axis, lateral axes immediately subtended on either side of the bud) no longer proliferating.</td>
</tr>
<tr>
<td>d_SCU</td>
<td>p₁</td>
<td>Distance to the nearest SCU on the same axis</td>
<td>The active growth of higher generation foliage along the same main axis inhibits bud release, and this effect is stronger the closer the growth is to the current bud.</td>
</tr>
<tr>
<td>d_stem</td>
<td>p₀</td>
<td>Distance to main stem</td>
<td>Hydraulic status lessens with distance to the base of the main branch stem.</td>
</tr>
<tr>
<td>r</td>
<td>r₀</td>
<td>Baseline bifurcation function</td>
<td>If all other parameters or variables are zero (or 1 for order), this is the rate of bifurcation. The effects of the variables are measured against this baseline.</td>
</tr>
<tr>
<td>o</td>
<td>rₜ</td>
<td>Shoot order</td>
<td>Apical control regulates bifurcation by shoot order. In the process of reiteration apical control is established; the main stem of SCUs are considered order 1.</td>
</tr>
<tr>
<td>g</td>
<td>rₜₑₚₑ</td>
<td>Shoot generation</td>
<td>A severe hydraulic constriction occurs at a junction formed by bud release from suppression, and this constriction worsens for each new generation.</td>
</tr>
<tr>
<td>n_over</td>
<td>rₜₑₚₑ</td>
<td>No. of overlapping foliated shoots</td>
<td>Foliage display is approximated by a count of the number of foliated shoots of the same generation or higher that intersect with the current shoot.</td>
</tr>
<tr>
<td>b_base</td>
<td>r_base</td>
<td>Bud age at SCU base when it was released from suppression</td>
<td>The hydraulic constriction caused by the junction formed by bud release from suppression lessens with distance along the main axis to the base of the axis.</td>
</tr>
<tr>
<td>d_base</td>
<td>None</td>
<td>Distance to SCU base</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. Optimization criteria

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Target</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>nfoliated</td>
<td>Quantitative domain</td>
<td>Number of foliated shoots</td>
<td>2500, 6500</td>
</tr>
<tr>
<td>nSCUs</td>
<td>Quantitative domain</td>
<td>Number of foliated SCUs</td>
<td>20, 50</td>
</tr>
<tr>
<td>length</td>
<td>Quantitative domain</td>
<td>Branch length (cm)</td>
<td>240, 430</td>
</tr>
<tr>
<td>arch_mod</td>
<td>Quantitative domain</td>
<td>Relative mechanical requirements for a given distribution of foliage</td>
<td>1: (R₃ &gt; R₂ &gt; R₁) : otherwise</td>
</tr>
<tr>
<td>μ_junctions</td>
<td>Prediction/alternative structure</td>
<td>Minimize</td>
<td>1/(n term) ∑₁₈ n term(i)</td>
</tr>
<tr>
<td>μ_path</td>
<td>Prediction/alternative structure</td>
<td>Minimize</td>
<td>1/(n term) ∑₈ n path(i)</td>
</tr>
<tr>
<td>load</td>
<td>Prediction</td>
<td>Minimize</td>
<td>√(P₁P₂) + √(P₃P₄) 27P₁ nfoliated</td>
</tr>
<tr>
<td>μ_over</td>
<td>Prediction</td>
<td>Minimize</td>
<td>1/(n term) ∑₈ n term(i)</td>
</tr>
</tbody>
</table>

Target ranges are set for each of the empirical criteria for the quantitative domain of model predictions (taken from Ishii and Ford, 2001), whereas the criteria for model prediction are minimized and two of which are used in analysis of alternative theories. RB defines the mean number of foliated shoots for a given shoot order (order 1, order 2, order 3). The prediction criteria are the mean number of cumulative junctions (μ_junctions), the mean path length from the base of the branch to active terminal shoots (μ_path), the relative mechanical requirements for a given distribution of foliage (load) and the mean number of overlapping foliated shoots (μ_over). Adapted from Kennedy et al. (2010).
(ii) Examination of features explicitly predicted by the model. Metrics used as criteria representing two causal processes were now included in the assessment: for water conduction and carbon balance. Two possible processes were considered for water conduction: that resistance to water flow increases with hydraulic path length (Zimmermann, 1978; Ewers and Zimmerman, 1984a, b); and that resistance increases as the numbers of branch junctions that water must flow through increase (Schulte and Brooks, 2003; Kennedy, 2008). Two possible processes were considered for carbon costs: that wood increment requirement depends upon mechanical stress (Pearcy et al., 2005) and this is likely to be less when foliage is consistently regenerated close to the main branch axis; and that foliage may develop opportunistically in places where there is less shading from neighbouring foliage. These processes and their interactions with epicormic reiteration are discussed by Ishii et al. (2007). Consequently, four predictive criteria defining requirements of branch design were used in assessment: (1) minimize the number of turns in the branching pathway; (2) minimize hydraulic path length; (3) minimize weight load on the branch; and (4) minimize foliage overlap. There are limits to such criteria because they are surrogates for more complete descriptions of the eco-physiological processes involved.

(iii) Tests between alternative structures in model formulation. Alternatives were examined in the morphological structures that may control water flow, i.e. path length and number of branch junctions at epicormic reiterations.

Stage 2. Evolutionary computation.

(a) Calculate the Pareto Optimal Set. The Pareto Optimal Set was calculated (Kennedy et al., 2010) using software developed by Reynolds and Ford (1999).

(b) Examine the adequacy of parameter search ranges. Exploration of the Pareto Optimal Set for adequacy of parameter range searches is described in Kennedy (2008) and Kennedy et al. (2009, appendix C).

(c) Examine if there are deficiencies in the mathematical structure of the model. Examination of previous versions of the model (Kennedy et al., 2004; Kennedy, 2008) established the effectiveness of the basic mathematical structure of the model.

Stage 3. Analysis of Pareto Set

(a) Examine the structure of the Pareto Optimal Set. A considerable number of parameterizations were obtained that achieved all empirical assessment criteria, and some that did not achieve all empirical assessment criteria yet achieved better values for subsets of the predictive criteria. The major result was that two distinct groups, non-dominated solutions were obtained for the four predictive criteria (Table 3). Those achieving low values for path length tended to achieve all of the empirical assessment criteria, referred to as Solution 1, and those achieving lowest values for the number of junction constrictions tended to achieve only two of the empirical criteria (branch length and number of foliated shoots), referred to as Solution 2. These two groups of solutions are non-dominated and there is no single parameterization that achieves a minimum in all four of the predictive criteria.

If we consider only non-dominated solutions they fall along a Pareto Frontier for mean hydraulic path length and mean number of hydraulic constrictions (Fig. 5A). The two groups also simulated different morphologies and had different values for certain parameters (example shown in Fig. 5B). The group in which the lowest values for path length were achieved appeared to have a form similar to old-growth Douglas-fir with groups of different epicormic generations arrayed along a main branch axis, and solutions in this group were able to satisfy all four empirical criteria. This is corroborated by the superior performance of this group with respect to the empirical criteria. The group in which the lowest values for the number of junctions were achieved appeared to have a form similar to an Abies species with a phalanx structure of foliage all of the same generation, and solutions in this group were unable to satisfy the empirical criteria corresponding to the number of SCUs and the expected architectural model (strictly decreasing bifurcation with increasing shoot order).

(b) Investigate distribution of parameter values between and within parameterizations of the different member groups of the Pareto Optimal Set. Inferences about how these distinct morphologies are achieved can be made by contrasting the parameter values found in each group through inspection of the Pareto Optimal Set. For example, parameterizations in Solution 1 (resembling Douglas-fir) tend to have lower values for order-1 bifurcation and a smaller restriction on the bifurcation of new epicormic shoots (Fig. 5B). This pattern of branch morphogenesis prevents simulated branches from achieving low values for the number of turns. Branches of Solution 2 tend to have much higher values for order-1 bifurcation rate and a more severe restriction on the bifurcation of new epicormic shoots. Those branches also tended to form SCUs well below the range specified for the SCU criterion and these relationships imply that, in the absence of extensive epicormic and SCU production, a branch

<table>
<thead>
<tr>
<th>Solution</th>
<th>Junction constrictions</th>
<th>Path length</th>
<th>Mechanical load</th>
<th>Foliage overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution 1</td>
<td>4.51</td>
<td>154</td>
<td>0.78</td>
<td>7.68</td>
</tr>
<tr>
<td>Solution 2</td>
<td>2.67</td>
<td>315</td>
<td>0.85</td>
<td>10.4</td>
</tr>
<tr>
<td>Solution 3</td>
<td>5.89</td>
<td>240</td>
<td>0.82</td>
<td>9.57</td>
</tr>
<tr>
<td>Solution 4</td>
<td>6.92</td>
<td>280</td>
<td>0.95</td>
<td>11.5</td>
</tr>
<tr>
<td>Solution 5</td>
<td>5.11</td>
<td>350</td>
<td>0.90</td>
<td>20.4</td>
</tr>
</tbody>
</table>

Solutions 1 and 2 are the only solutions among the five that are mutually co-dominant and not dominated by other solutions and represent distinct groups of parameterizations found repeatedly. Solution 1 performs better for path length while Solution 2 performs better for junction constrictions. Solutions 3, 4 and 5 all perform worse than Solutions 1 and 2 for at least one criterion and perform better in none.
must more rigorously maintain its major axes. This pattern of branch morphogenesis results in branches with relatively longer path lengths and fewer turns.

Stage 4. Define model uncertainty.

The Pareto Optimal Set and important aspects of variation in parameter values are described in Stage 3. A summary of uncertainty in the model is given in Table 4, including the important components of criteria that would provide a more effective assessment of the model and would require additional data. A limitation is that visual comparisons of plant form were used. Success was attributed to the model in producing SCUs ‘close’ to the main branch axis but this should be defined quantitatively particularly because of the

---

**FIG. 5.** (A) Pareto Optimal Frontier for the criteria that are optimized, showing the trade-off between path length and number of turns. (B) Corresponding Pareto Optimal Set for one of the model parameters, the reduction in bifurcation for new epicormic shoots. There is a vertical line at −150. Parameter values greater than −150 occur for Douglas-fir-like branches; those less than −150 occur for *Abies*-like branches. The inserts provide a view of simulated branches with cm scales indicating distances from the point of origin of the branch and colours the age of foliage in successive generations of epicormic branching.
posited importance of hydraulic path length. The criteria designed for analysis of causal processes could be improved, for example if data were collected on changes in resistance to water flow within the branching network. However, the details of the proposed causal processes may require development of another FSPM to examine these details. Although BRANCHPRO3 establishes that these processes may contribute to branch longevity, the large size of the model may make it unsuitable for more detailed modelling. The ability of the model to predict two types of branching when different sets of parameter values were found and alternative criteria were satisfied (hydraulic path length vs. hydraulic junctions) reduces some uncertainty. This is because the architectural theory on which the model is based can now be considered to have greater explanatory power. In this sense the model is more completely specified. However, only indirect measures were used in assessment. It is important to note that the simulated branches that reduced numbers of branch junctions appeared more like Abies than Pseudotsuga and did not satisfy criteria for the quantitative domains. Those domains would have to be defined specifically for Abies through collection of data.

**DISCUSSION**

FSPMs synthesize information about plant function and structure. Typically they define aspects of the dynamics of plant growth, i.e. how processes change in response to the environment or developments in plants themselves. The dynamic process defined by BRANCHPRO3 is an example of structural homeostasis in the continued regeneration of foliage in response to changes caused by growth itself and the internal environment of the plant. Through use of the model it was possible to define limits to how long this homeostasis could continue (Kennedy, 2010; Kennedy et al., 2010).

Many FSPMs are used in progressive analysis of plant systems. They are developed over time by changing their structure to achieve more effective explanations. Assessment of FSPMs should complement this development by providing information on how improvements might be made. The objective is to analyse whether the components of an FSPM are described and organized effectively in the model and this requires assessment against a number of different criteria. There are two essential interacting features of the illustrated assessment system: use of evolutionary computation to explore the parameter space of the model and selection of effective criteria.

Evolutionary computation explores the possible sets of parameterizations that can achieve a set of criteria. The way in which the model functions can be explored by widening or narrowing allowed ranges of parameters and repeating the evolutionary computation. Similarly, when information about assessment criteria is uncertain their ranges can be manipulated, also revealing aspects of how a model functions: Turley and Ford (2009) provide an example. This use of evolutionary computation analysis provides an advantage over sensitivity analysis because it can be directed to explore within known domains of realistic outputs while sensitivity analysis explores what might be possible with particular combinations of parameter values without restriction on feasibility of outputs.

An important feature of multi-criteria assessment is the selection of criteria. The categories used in the example illustrate the types of criteria that might be valuable in developing effective explanations but there are two important challenges. The first is in quantifying the description of plant form. Visual interpretation was used in this analysis as it is in many FSPM analyses but it remains a weakness – the human eye can be deceived. Rather than seeking a quantitative description of the integrated plant body one approach is to develop specific
questions about plant form in relation to model functioning, e.g. in the example used here distance of SCUs from the main branch axis would be quantifiable.

The second challenge is to align assessment criteria with particular aspects of model function and to ensure, where possible, that criteria are selected to sense across the range of model functioning. There can of course be difficulties in this. We often deliberately model what we cannot easily measure at the scale of our investigations. In the example used here the explanation provided by the model is implied to be in water relationships and self-shading of foliage, neither of which were measured but indicate a further research direction.

It is important to draw a distinction between assessment of model structure and statistical inference. Methods of statistical inference assume that the structure of the statistical model is correct and that there is a sufficient statistic for testing the fit of a set of data to the model (Fisher, 1922). A sufficient statistic exists when no other statistic can be calculated that provides additional information about the model parameters. When attempts are made to use methods of statistical inference to assess process models, a particular measure may be chosen without consideration as to whether it is a sufficient statistic or not. In statistical inference parameters are estimated to give some measure of goodness of fit. For models with many parameters, some reasonable fit can usually be found because the model accommodates to the data through variation in parameter values and in practice many sets of parameter values might do that equally well (e.g. Beven, 2006). In statistical inference, attempts are made to reduce the problem of accommodation caused by model complexity by applying the principle of parsimony and statistics such as the Bayesian Information Criterion (Schwarz, 1978) or Akaike Information Criterion (Rao and Wu, 2001) are used that penalize models with more parameters. In contrast, the objective with FSPMs is not to find the most-parsimonious model that fits one statistic to some data. Rather, it is to assess the structure of the botanical and environmental model. Are the right components included and in the right way and does the model successfully predict all that it should? For this purpose we need to use multiple criteria.

The definition of uncertainty provided by the approach described here is analytical in dealing with different components of the defined explanation. This is particularly appropriate for FSPMs, where the scientific goal is to advance botanical knowledge through the use of simulation models. Other fields of research also use the concept of uncertainty in an analytical framework, for example the analysis of measurements in physics (Ignacio, 2002) and the analysis of hydrology models in which evolutionary computation and Pareto concepts are employed (Efstathiadis and Koutsoyiannis, 2010).

ACKNOWLEDGEMENTS

We thank Shawn Behling, Mandi Fix and Bert Loosmore for helpful comments on the manuscript and are grateful to Genevieve Hines and an anonymous reviewer for their constructive critiques.

LITERATURE CITED


